Sensing Geographical Impact Factor of Multimedia News Events for Localized Retrieval and News Filtering

Xu Zhang[1,2], Jin-Tao Li[1], Yong-Dong Zhang[1], and Shi-Yong Neo[3]

[1] Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[2] Graduate School of the Chinese Academy of Sciences, Beijing 100039, China
[3] National University of Singapore, 3 Science Dr, Singapore 117543
{zhangxu,jtli,zhyd}@ict.ac.cn,
neoshiyo@comp.nus.edu.sg

Abstract. News materials are reports on events occurring in a given time and location. Looking at the influence of individual event, an event that has news reported worldwide is strategically more important than one that is only covered by local news agencies. In fact, news coverage of an event can accurately determine the event's importance and potential impact on the society. In this paper, we present a framework which extracts the latent impact factor of events from multimedia news resources by a geographical approach to support: (a) localized retrieval for end users; and (b) pre-screening of potential news elements that should be filtered for use by web monitoring agencies.

Keywords: Multimedia News Impact Factor, Retrieval, Filtering.

1 Introduction

Information retrieval especially in the domain of multimedia news is increasingly important with the ever increasing amount of multimedia data. With the rise of multimedia news search engines such as Yahoo!, Google and MSN, it has become extremely easy for end-users to gain access to the wide range of published multimedia news available on the Web. This has however created a scenario where the end-user is subjected to an extensive amount of unrelated and undesired information. In particular, regulating agencies are also facing difficulties in filtering undesirable news materials, especially those in non-text formats such as videos. Consequently, to handle this problem and enhance news video retrieval, some form of advanced video processing is needed.

Most prior researches in video processing rely strongly on features within video source only for retrieval [1], [2]. A recent approach [3] demonstrated that features from external information such as blogs and online news articles may also be useful in improving retrieval. In this paper, we propose another useful facet of news, the geographical news coverage, or what we term as the event impact factor. News reports consist of essential events occurring in a given time and location, with different news agencies having its own degree of involvement. A piece of news that has a worldwide coverage is termed more strategically important than one that is only covered by local news

S. Boll et al. (Eds.): MMM 2010, LNCS 5916, pp. 567–576, 2010.

agencies. In fact, the news coverage of an event has direct implications to its impact on the society. Based on this observation, we propose a framework to extract and utilize the derived event impact factor as a major enhancement for: (a) supporting localized multimedia news retrieval; (b) acting as a gauge for filtering of sensitive news for data governing agencies.

Sensing the Impact factor. The impact factor of an event can be inferred from the news coverage of the event. In reality, a unique event can have different video presentations and be narrated in different languages for different locations or lingual groups. In order to determine the news coverage for an event, it is necessary to find all the related news materials reporting about the event. We therefore perform clustering of the collected news materials to obtain coherent groups of news for different events. The process is similar to topic detection and tracking (TDT) [4]. Using the groups obtained, the news coverage of an event can then be inferred through the strategic influence of each news agency. For example: "[Straits Times, Channel 5] Singapore", "[人民日报(People's Daily), CCTV] China", "[CNN, NYT, MSNBC] US". Subsequently, the impact factor matrix is derived for each cluster of news materials to support retrieval and filtering.

Utilizing the Impact factor for localized retrieval and news filtering. To leverage the impact factor matrix, we first need to know an event's actual occurrence location. We determine this by space and frequency analysis of location entities mentioned in news articles and news videos. There might be more than one location entity mentioned, but through the analysis, confidence score can be assigned to each entity. One useful heuristic rule to aid in confidence scoring is that news reports will commonly begin by highlighting the event location. At present, news search engines localize searches through the detection of IP address. Our proposed impact factor can similarly be exploited like IP address to further rank news accordingly. The rationale is that impact factor scores can effectively depict an event's importance. In retrieval we can allow users to indicate several locations which they are interested so as to view the highly important news in these locations. As for news filtering, we can forbid users from a targeted location "T" in gaining access to sensitive news materials, for example politics-related news videos. As an illustration, a particular event E which occurs in "T" may have a large amount of news coverage in many countries (high impact factor), but significantly low impact factor in "T" itself. This scenario can signal that news regarding the event could be sensitive in nature for "T". The intuition for this is built upon a reverse psychologically effect where local media in "T" have likely done preventive measures to block sensitive news.

2 Sensing Event Impact Factor

The influence of an event is location sensitive, as people living in different parts of the world are likely to see only news which is made known to them through their local media broadcasters. Thus, it makes perfect sense to rank news according to the location sensitive impact factor. Figure 1 shows the overall framework for event impact factor sensing. The first step to understand the news coverage is to obtain semantic groups of news reports related to an event. News reports of the same event can be in the form of video, text or audio, and presented differently across multiple agencies even in different languages. It is necessary to pre-cluster related news materials on a per-event basis. Catering to this, we construct an event space to model events and utilize an unsupervised temporal multi-stage clustering algorithm to obtain coherent groups of news. With this grouping, it is possible to infer the coverage of an event and its impact by analyzing the strategic influence of the news agencies reporting the event.

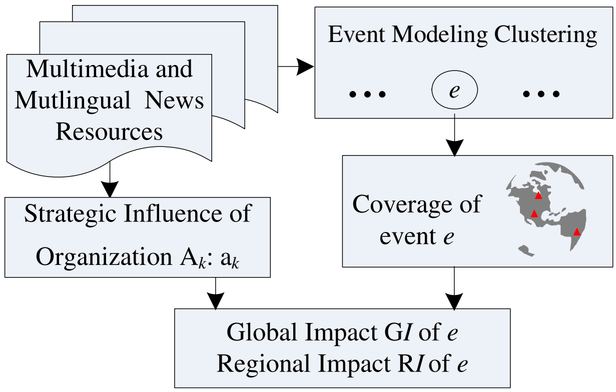

Fig. 1. Framework for Sensing Impact Factor

2.1 Multimedia News Event Modeling and Clustering on News Video and Online Articles

There are subtle differences between news and events. While an event is the occurrence of an incident, news is the report on a series of events. Regardless of the media format in publication for a news report, it will contain various event aspects such as: *Location*, *Time*, *Subject*, *Object*, *Quantity*, *Action* and *Description*, etc. Thus by translating a piece of news onto a distinct point in a multi-dimensional event space, where each dimension is an event aspect, we can effectively use news materials to model events. For this task, we have chosen the event feature space and methods according to the observed effectiveness in [3]. The features used in our event space consist of: (a) text entities from speech; (b) high level features (HLF); (c) near duplicate (ND) information. They are extracted from videos and online news articles. In our model, we define essential text event entities as follows: **Location** {country, city, county, places of interest, etc}, **Time** {video timestamp or specific date mentioned, etc}, **Object** {tangible like car, people, intangible like war, oil prices}, **Subject** {person's name, organization, etc}, **Quantity** {numerical}, **Action** {death, birth, murder, etc} and **Descriptions** {other deeds}.

Apart from text entities, HLF is also utilized, which are extracted from the visual content of videos through support vector machine analysis on low level features. As shown in [5], the use of HLF allows news to have more semantic information. Coupled with this, our model is further enhanced by ND information [6], which can easily provide the linkage between different news video events. This is because reports of the same event can sometimes have the same footage or slightly different footages.

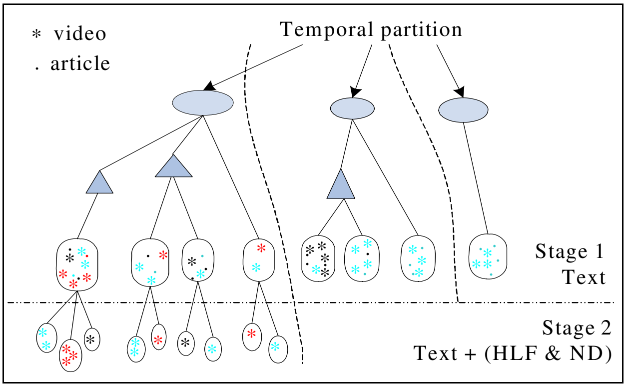

Fig. 2. Temporal Multi-stage Event Clustering

In order to obtain coherent groups of news from news video and online articles, an unsupervised temporal multi-stage clustering algorithm as shown in Figure 2 is used. The multi-stage clustering framework uses text during the first stage of clustering and a combination of text, HLF and ND for the second stage. The aim of the first stage is to identify possible events using text, which is one of the best feature for event detection and tracking. However, text from speech transcripts may be erroneous and insufficient. Therefore we supplement it with text from parallel news articles. This combination enables key entities in events (from text) to be correctly extracted for effective event clustering. In the second stage, HLF and visual features are utilized to refine the initial clusters. News resources are then divided into temporal partitions before clustering. The aim of temporal partitions is to reduce the computation time, and provide a smaller clustering space from which better clustering results can be obtained. More details can be obtained from [3].

2.2 Strategic Influence of News Agencies and Event Impact Factor

As news materials are copyrighted, details of publishing agencies are usually appended with the news elements. Considering such a characteristic, we introduce a two-stage approach to map events into their news coverage areas with respect to agencies by: a) identify the agencies and their strategic area of influence; and b) map the news reports or videos in clusters from Section 2.1 to corresponding agencies and tabulate their impact scores. We use a list containing over 1000 news agencies for computing the impact scores. This list is a subset of which Google online news [7] performs their crawling. The approximate breakdown for the new agencies across countries is given in Figure 3.

In many countries, events that are reported by authoritative agencies are relatively more influential than those reported by smaller agencies. The strategic penetration depth (or circulation) of the agency also plays a significant role in determining the amount of potential readers and viewers. This strategic reach of an agency can be

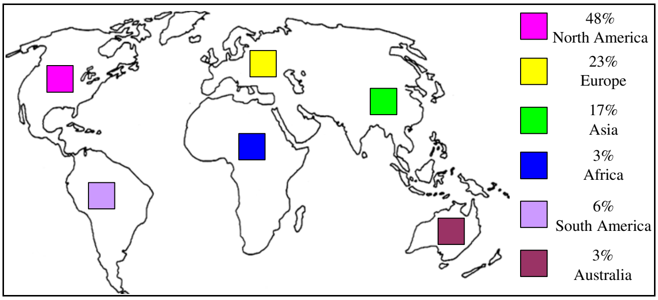

Fig. 3. Strategic breakdown of news agencies at marco-continental level

estimated through its local viewer/readership population (at county, province or country level, depending on the area of circulation). Another observation about authoritative agencies is that they tend to produce larger amount of news reports. It is therefore possible to evaluate the agency strategic importance (ASI) by counting the daily reports. We approximate this value by the number of reports of each news agency during a certain period. The ASI a_k of each news agency A_k is defined in equation 1 to provide the underlying strategic importance.

$$a_k = n_{A_k} \cdot p_{A_k} \tag{1}$$

where n_{A_k} is the total number of reports issued by A_k during the given period and normalized to [0,1] by standard normalization; p_{A_k} is the population normalized to [0,1] in the location of A_k which is extracted together with agencies.

With the individual ASI a_k, it is then possible to measure the impact of an event appropriately with respect to various locations. In particular, we consider the impact factor of an event *e* at an individual country-level since this is most appropriate for many applications. The regional impact factor $RI_{e|L}$ of *e* at location *L* is given in equation 2.

$$RI_{e|L} = \lambda \sum_i (a_i \cdot n_e) + (1-\lambda) \sum_j (a_j \cdot len_e) \tag{2}$$

where n_e is the total of text articles about event *e* (for text-based news) and len_e is the length of videos about event *e* (for video news), both of which are normalized to [0,1] by standard normalization. Since video news usually emphasizes events that are most interesting or important due to expensive and limited airtime, we give higher weight to video news. Empirically the weight λ for text news is set as 0.3, weight for video news as 0.7. The impact of an event to various countries is tabulated in the form of an impact factor table. In addition, a global impact factor GI_e of *e* can be calculated by aggregating the scores of all reporting locations. Figure 4 shows the impact of event "Sichuan Earthquake" collated from news report during the period 12th to 22nd of May2008 normalize to 4 shading intensity with the region of highest impact in black.

Fig. 4. Impact of event "Sichuan Earthquake"

3 Utilizing Event Impact Factor

After getting the impact factor table, a prior step to extract the event occurrence location is carried out by spatial and frequency analysis of location text entities from each news cluster. With the available event feature space, the features are then employed for: (a) localized news retrieval; and (b) pre-screening of potential news elements that should be filtered for data monitoring agencies.

3.1 Getting the Event's Occurrence Location

Content of news usually contains more than one location entity, which makes it challenging to determine the event's occurrence location. We make decisions with the help of text pragmatics and video cinematic in a rule-based fashion considering the following two characteristics. First, it is observed that news is often presented in a top-down fashion where highlights of news are usually presented at the beginning. This characteristic increases the likelihood that the actual event location appears in the beginning rather than at the end. Second, the actual location entity is likely to be mentioned more frequently than other location entities for emphasis.

Combining the above two characteristics, we propose a *spatial-freq* approach to determine event's location of occurrence. Let $LE=\{L_1,\ldots, L_i,\ldots L_n\}$be all location entities in a news article or video, the event's occurrence location L_e is defined as follows:

$$L_e = \arg\max_{L_i}\left(\beta \cdot r_{L_i} + (1-\beta)\cdot f_{L_i}\right) \tag{3}$$

where r_{L_i} is the reciprocal of number of words before the first appearance of location entity L_i, and f_{L_i} is the frequency of L_i in news content; β is set as 0.6 to emphasize the relative position of L_i. Recognizing the first appearance of L_i requires high complexity in checking for each word. Thus, we use Aho-Corasick algorithm[8] which is designed for fast matching of strings and finding the first possible substrings for our problem.

3.2 Localizing Geo-based News Retrieval

Current online news search engines facilitate localized news searching through users' IP addresses and display news from their localized news agencies. (For example:

CNN and NYT for US, "Straits Times" for Singapore, etc) This rationale is straightforward as news agencies situated near the location of the user are likely to deliver news which is of interest to that user. However, if there is no news coverage of an event by that local broadcaster, it is possible for users to miss that event. With the knowledge of an event location together with its global and regional impact factors, our system aims to further enhance localized retrieval with respect to user's interest, location and IP. Given an event e with its occurrence location L_e, global impact GI_e and regional impact $RI_{e|L}$, the modified ranking score $\bar{s}_{R_{e|L}}$ for news report R_e (article or video) about e at location L is shown in equation 4.

$$\bar{s}_{R_{e|L}} = \theta \cdot \left(\frac{RI_{e|L}}{GI_e} \right) + (1-\theta) \cdot s_{R_e} \tag{4}$$

where s_{R_e} is the original ranking scores generated by initial retrieval algorithm on a scale from 0 to 1 (1 is most relevant) for each news report, the weight θ for impact-based ranking score is empirically set as 0.4 for aggregation.

3.3 News Filtering for Network Monitoring

The idea of online news monitoring is to restrict users within a location to news materials which could be sensitive in nature, as monitoring organizations do not have access control over overseas news agencies. Traditional news filtering systems performs filtering using a predefined list of key words and various matching strategies. That approach can only detect sensitive text in news according to a preference list, and have much limitation on news video. With the discovered event impact factor, it is now possible to estimate the sensitivity $S_{e|L}$ of an event e with respect to location L as in Equation 5.

$$S_{e|L} = \frac{GI_e - RI_{e|L}}{GI_e} \tag{5}$$

where the regional impact $RI_{e|L}$ of e in L and its global impact GI_e can be accessed through the impact factor table generated in Section 2. The larger the value of $S_{e|L}$ is, the event e is more sensitive in the location of L. This criterion is based on the assumption that events should be reported more in its happening places.

4 Experiments

In this section, we will present the experimental results to the proposed tasks of retrieval and filtering. We leverage news video corpus from TRECVID 2006[5] testing dataset, consisting of about 160 hours of multilingual news videos in English, Arabic and Chinese. In addition, we crawled 463,000 news articles through various online news archives, including the Highbeam Research database [9] which archived over 1000 sources. These archived news articles are reported during the same period with TRECVID news video.

4.1 Clustering Performance

Clustering performance is crucial as it will have a predecessor effect on the later applications. To evaluate the clustering performance, manual assessment is done to measure the clustering quality using human annotators. As it is too labor-intensive to screen through all clusters, a subset is selectively chosen for evaluation. This subset stretches over a 7-day period equivalent to approximately 15% of the testing corpus. During evaluation, the human annotators are asked to access inter-cluster and intra-cluster correctness. Three runs are designed to test the effectiveness of: (1) text features, (2) high level features and (3) near duplicate information. The results are tabulated in Table 1.

Table 1. Performance of clustering

TRECVID2006	Baseline	A	B	C
Prec.	0.376	0.418	0.437	0.478

where Baseline: (text event entities with temporal partitions)
A) baseline+ HLF
B) baseline + near duplicate
C) baseline+ HLF + near duplicate

From Table 1, we can draw the following conclusions. First, improvements can be seen with the addition of HLF and near duplicated information. In particular, near duplicate information seem more effective. This can be attributed to the fact that an event can have similar footages across multiple news report and this makes near duplicate information important. The run which utilizes both HLF and near duplicate information yields the best performance, demonstrating the complementary boosting effect.

4.2 Effects of Localizing Geo-based Retrieval

This series of experiment is designed to investigate the effects brought by the enhancement of localization. Ten users with prior knowledge of news video retrieval are selected for the experiments. The users are asked to use the retrieval system and evaluate the retrieved results on a scale of 1 to 5 (5 as with best performance). The following are two sample questions from the evaluation questions list: *How relevant is the news to the input query? Does the news reflect the location you are interested?*

For comparison, we make use of the previous news video system as in [3]. The results are tabulated in Table 2, which shows that users prefer our system with the localization effect. Interactive user interface is suggested by users as part of our future system enhancement.

Table 2. User rating (average)

	[3] without impact factor	[3] with impact factor
Score	3.97	4.64

4.3 Sensitive News Video Filtering

The last series of test is designed to access the effectiveness of the system in ranking sensitive news video. For this test, we select 5 locations of interest: China, India, Saudi Arabia, Russia, and Japan. For each location, the system ranks news videos according to the sensitivity score S generated by equation (5). Top n (Max: n= 100) News videos with $S_{e|L}$ above threshold T is selected as potentially to be filtered, where T is determined by Gaussian estimation as $\overline{S}_{e|L} + 3\delta_s$, where $\overline{S}_{e|L}$ is mean of $S_{e|L}$, $\delta_{S_{e|L}}$ is variance of $S_{e|L}$. As there is no ground truth available to access the performance, we employ 5 human assessors to manually judge the sensitivity. Each human assessor will view the news video clip of the top n list and evaluate the sensitivity by three categories: *sensitive*, *not sensitive*, and *unable to judge*; depending on the impact on politics, military or economic at that location. The purpose of this manual assessment is to compare between the judgments of sensitive news videos from our system and from human perception. Subsequently, the agreement values are measured in term of *Fleiss' kappa* for measuring agreements[10].

Table 3. Kappa User Agreement (k)

TRECVID2006	Highest	Lowest	Average
Inter User	0.89	0.51	0.74
User vs. System	0.78	0.46	0.63

From Table 3, we can conjecture that the agreement between users is quite high (an average of 0.74). This is also true for the agreement score between user and our system on sensitive news (an average of 0.63). Although little news which is not sensitive to users is potentially filtered by our system, this is much because of the threshold is too strict. Better estimation of threshold can improve the results. Overall, this signifies that our system is able to filter and detect sensitive news quite reasonably.

5 Conclusion

In this paper, we presented the idea of mining geographical impact factor of news events to support news video retrieval. Two applications are proposed based on our framework: (a) localized news retrieval for end users; and (b) pre-screening of potential news elements for data monitoring agencies. Preliminary results using the TRECVID 2006 dataset and archived news articles demonstrate the effectiveness and usability of our framework. For future work, we are looking into analyzing different opinions across multiple news agencies to summarize and provide more various aspects of news events.

Acknowledgements

This work was supported by National Basic Research Program of China (973 Program, 2007CB311100), National High Technology and Research Development

Program of China (863 Program, 2007AA01Z416), National Nature Science Foundation of China (60873165、60802028), Beijing New Star Project on Science & Technology (2007B071) and Co-building Program of Beijing Municipal Education Commission.

References

1. Chang, S.-F., Winston, J.W., Kennedy, L., Xu, D., Yanagawa, A., Zavesky, E.: Columbia University TRECVID-2006 Video Search and High-Level Feature Extraction. Gaithersburg (2006)
2. Campbell, M., Hauboldy, A., Ebadollahi, S., Joshi, D., Naphade, M.R., Natsev, A.P., Seidl, J., Smith, J.R., Scheinberg, K., Tešić, J., Xie, L.: IBM Research TRECVID-2006 Video Retrieval System
3. Neo, S.-Y., Ran, Y., Goh, H.-K., Zheng, Y., Chua, T.-S., Li, J.: The use of topic evolution to help users browse and find answers in news video corpus. In: Proceedings of the 15th international conference on Multimedia (2007)
4. James, A., Ron, P., Victor, L.: On-line new event detection and tracking. In: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval (1998)
5. `http://www-nlpir.nist.gov/projects/trecvid/`
6. Zheng, Y.-T., Neo, S.-Y., Chua, T.-S., Tian, Q.: The use of temporal, semantic and visual partitioning model for efficient near-duplicate keyframe detection in large scale news corpus. In: Proceedings of the 6th ACM international conference on Image and video retrieval (2007)
7. `http://news.google.com/`
8. Aho, A.V., Corasick, M.J.: Efficient string matching: an aid to bibliographic search. Commun. ACM 18(6), 333–340 (1975)
9. `http://www.highbeam.com/web/`
10. Landis, J., Koch, G.: The measurement of observer agreement for categorical data. Biometrics (1977)